Research Paper ■

The Determination of Relevant Goals and Criteria Used to Select an Automated Patient Care Information System:

A Delphi Approach

JOAN K. CHOCHOLIK, MHA, SUSAN E. BOUCHARD, JOSEPH K. H. TAN, PhD, DAVID N. OSTROW, MD, MA

A b s t r a c t   **Objectives:** To determine the relevant weighted goals and criteria for use in the selection of an automated patient care information system (PCIS) using a modified Delphi technique to achieve consensus.

Design: A three-phase, six-round modified Delphi process was implemented by a ten-member PCIS selection task force. The first phase consisted of an exploratory round. It was followed by the second phase, of two rounds, to determine the selection goals and finally the third phase, of three rounds, to finalize the selection criteria.

Results: Consensus on the goals and criteria for selecting a PCIS was measured during the Delphi process by reviewing the mean and standard deviation of the previous round's responses. After the study was completed, the results were analyzed using a limits-of-agreement indicator that showed strong agreement of each individual's responses between each of the goal determination rounds. Further analysis for variability in the group's response showed a significant movement to consensus after the first goal-determination iteration, with consensus reached on all goals by the end of the second iteration.

Conclusion: The results indicated that the relevant weighted goals and criteria used to make the final decision for an automated PCIS were developed as a result of strong agreement among members of the PCIS selection task force. It is therefore recognized that the use of the Delphi process was beneficial in achieving consensus among clinical and nonclinical members in a relatively short time while avoiding a decision based on political biases and the "groupthink" of traditional committee meetings. The results suggest that improvements could be made in lessening the number of rounds by having information available through side conversations, by having other statistical indicators besides the mean and standard deviation available between rounds, and by having a content expert address questions between rounds.

■ **JAMIA.** 1999;6:219–233.

Affiliations of the authors: Vancouver Hospital and Health Sciences Centre, Vancouver, British Columbia, Canada (JKC, DNO); Snohomish, Washington (SEB); University of British Columbia, Vancouver (JKHT).

Correspondence and reprints: J. K. Chocholik, MHA, Manager, User Readiness Program, Care 2000 Project, Vancouver Hospital, 855 West 12th Avenue, Vancouver, BC, Canada V5Z 1M9. e-mail: ⟨jchochol@vanhosp.bc.ca⟩.

A patient care information system (PCIS) selection task force was given the task of recommending to the board of the Vancouver Hospital and Health Sciences Centre (VHHSC) an integrated, computerized PCIS. The VHHSC is a multisite complex providing a range of tertiary-care services to the citizens of British Columbia, Canada. The complex has five Vancouver-based sites and three regional sites, with a total of 1,900 beds serviced by 1,500 physicians and approximately 9,500 caregiver and administrative staff. With

the exception of pediatric and obstetric services, a wide range of medical, surgical, and psychiatric services are provided, including highly specialized provincial programs for leukemia and bone marrow transplants, spinal cord injuries, burns, solid organ transplantation, and acute trauma services.

The PCIS selection task force was given a six-month mandate to evaluate two computerized PCIS systems and present a recommendation. Their key responsibilities during this project were to develop an in-depth knowledge of each product in relation to the clinical activities performed at the hospital, communicate their activities and decisions to their respective constituencies and VHHSC executive management, and make a recommendation to the board.
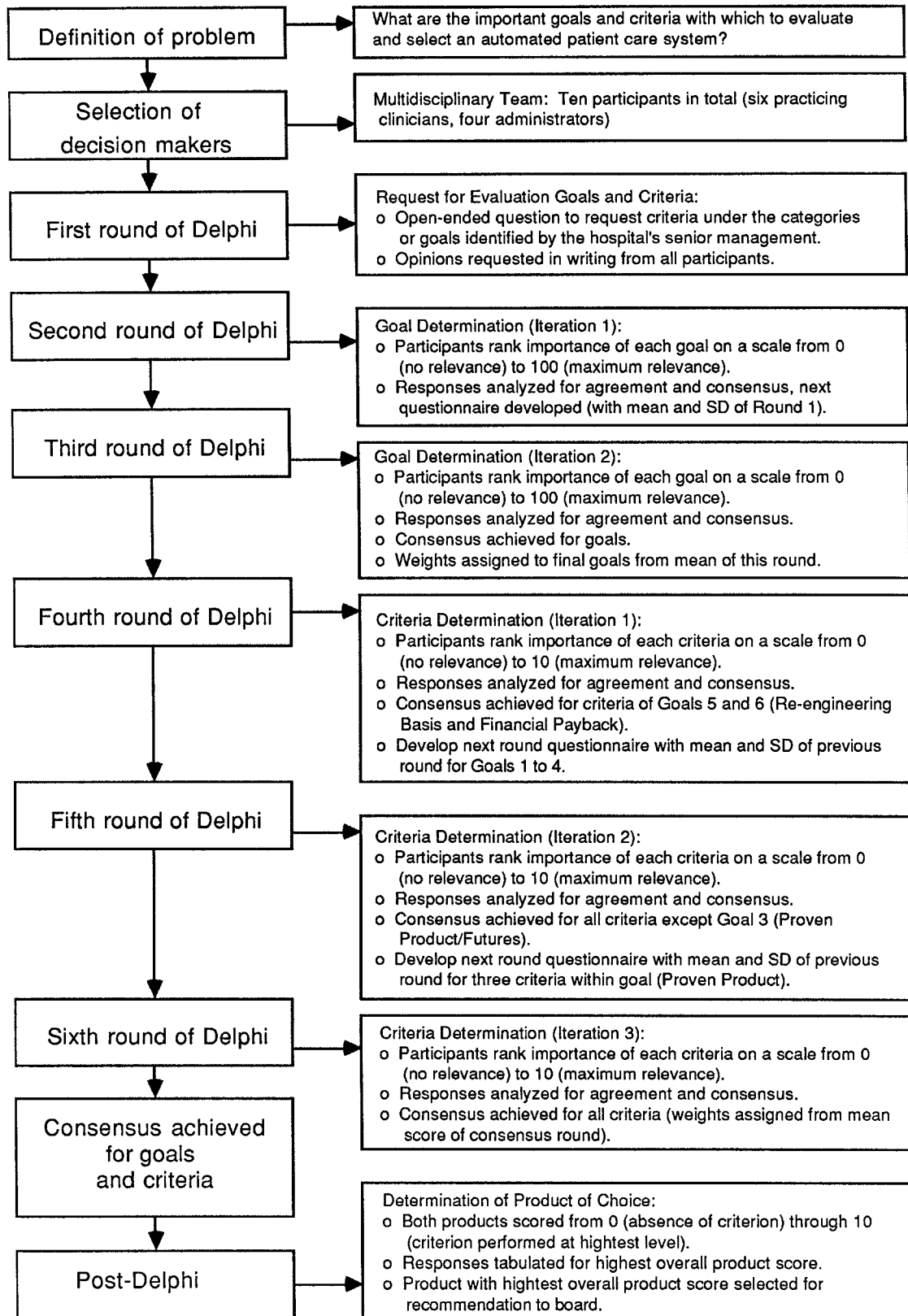
Prior to the formation of the PCIS selection task force, which included representatives from the clinical services of the hospital, there was a "preselection" phase managed by administration. Two major activities occurred during this phase. A review of the literature was performed by the hospital's information systems support staff, and a summary of the findings was made available to the PCIS selection task force panel members as part of their preparation during the actual selection project. In addition, the hospital retained the services of an international management consulting firm specializing in health care systems. Their mandate was to review the marketplace and recommend the two most suitable PCIS software products for implementation at an institution of the size and complexity of VHHSC. The acceptance of the recommendation of two products by executive management served as the starting point for the selection recommendation project managed by the PCIS selection task force.

The first part of this six-month selection recommendation project, known as the "pre-Delphi" phase, involved extensive fact finding on the two products through on-site product demonstrations, "hands-on" trials of clinical scenarios using the PCIS software products, visits to vendor headquarters, and client site visits. All PCIS selection task force members—in effect, the Delphi panel members—were required to participate actively by attending the on-site demonstrations, developing and leading the trials of the clinical scenarios using both software products, and visiting vendor headquarters as well as nine Canadian and American client sites, which were similar in size and complexity to VHHSC and had implemented one of the two PCIS software products under review. The panel members held a number of debriefing and data collection wrap-up meetings following these site vis-

its, during which each member would relay their findings to the rest of the group. The client site visits were an essential part of the process, in that they gave the panel members, particularly the clinicians, the opportunity to gain practical insights into both products as well as identify relevant criteria for selection of a PCIS product. In addition, the clinical members of the panel reviewed literature on the implementation of automated clinical systems and used the panel members from Information Systems as advisors to become knowledgeable about the factors relevant to the selection of software products. In effect, the first part of this selection project served as a means of educating the caregivers about what factors would be relevant in the selection of a PCIS about which they initially had little knowledge. Gaining this knowledge through first-hand experience was viewed as an advantage over having criteria supplied to the panel prior to the site visits. During the second part of the selection process, the four-week Delphi phase, this multidisciplinary task force used several consensus-building techniques to arrive at a recommendation—namely, brainstorming, the nominal group meeting technique, and a decision Delphi process.

The main objective of using the Delphi technique in the latter stage of the detailed selection process was to determine and weight the relevant selection goals and the specific criteria within each goal. These weighted goals and criteria were then used to evaluate two software products to determine a product of choice. This modified Delphi process consisted of six rounds (Figure 1). The outcome of the Delphi process was the achievement of consensus on the goals and criteria and their respective weights, each weight being equivalent to the mean score of the round when consensus was achieved. Using the weighted goals and criteria, the two software products were evaluated and the product with the highest overall product score was recommended for purchase and implementation.

It was the decision of the chair of the PCIS selection task force to use a Delphi process to arrive at the PCIS software product decision. The Delphi technique was chosen primarily because the PCIS software recommendation would affect the entire caregiver community and thus required a collective group opinion from people with diverse perspectives, experience, and expertise. Making a decision while avoiding the "groupthink," or bandwagon effect, of traditional committee meetings and the influence of more powerful committee members required a process that could address these concerns. Janis[1] describes "group-

| Definition of problem | → | What are the important goals and criteria with which to evaluate and select an automated patient care system? |

| Selection of decision makers | → | Multidisciplinary Team: Ten participants in total (six practicing clinicians, four administrators) |

| First round of Delphi | → | Request for Evaluation Goals and Criteria:
o Open-ended question to request criteria under the categories or goals identified by the hospital's senior management.
o Opinions requested in writing from all participants. |

| Second round of Delphi | → | Goal Determination (Iteration 1):
o Participants rank importance of each goal on a scale from 0 (no relevance) to 100 (maximum relevance).
o Responses analyzed for agreement and consensus, next questionnaire developed (with mean and SD of Round 1). |

| Third round of Delphi | → | Goal Determination (Iteration 2):
o Participants rank importance of each goal on a scale from 0 (no relevance) to 100 (maximum relevance).
o Responses analyzed for agreement and consensus.
o Consensus achieved for goals.
o Weights assigned to final goals from mean of this round. |

| Fourth round of Delphi | → | Criteria Determination (Iteration 1):
o Participants rank importance of each criteria on a scale from 0 (no relevance) to 10 (maximum relevance).
o Responses analyzed for agreement and consensus.
o Consensus achieved for criteria of Goals 5 and 6 (Re-engineering Basis and Financial Payback).
o Develop next round questionnaire with mean and SD of previous round for Goals 1 to 4. |

| Fifth round of Delphi | → | Criteria Determination (Iteration 2):
o Participants rank importance of each criteria on a scale from 0 (no relevance) to 10 (maximum relevance).
o Responses analyzed for agreement and consensus.
o Consensus achieved for all criteria except Goal 3 (Proven Product/Futures).
o Develop next round questionnaire with mean and SD of previous round for three criteria within goal (Proven Product). |

| Sixth round of Delphi | → | Criteria Determination (Iteration 3):
o Participants rank importance of each criteria on a scale from 0 (no relevance) to 10 (maximum relevance).
o Responses analyzed for agreement and consensus.
o Consensus achieved for all criteria (weights assigned from mean score of consensus round). |

| Consensus achieved for goals and criteria | | |

| Post-Delphi | → | Determination of Product of Choice:
o Both products scored from 0 (absence of criterion) through 10 (criterion performed at hightest level).
o Responses tabulated for highest overall product score.
o Product with hightest overall product score selected for recommendation to board. |

Figure 1 Delphi process for the PCIS Selection Project at Vancouver Hospital.

think'' as a "mode of thinking that people engage in when they are deeply involved in a cohesive in-group, when the members' strivings for unanimity override their motivation to realistically appraise alternative courses of action," which can result in a deterioration of mental efficiency, reality testing, and moral judgment. The chosen method was a modification of the conventional Delphi process.[2] Each of the participants knew the others. In addition, the panel was composed not just of experts in one field but of a mixture of specialists in clinical care, patient care administration, and information systems. The objective and framework was most similar to the decision Delphi[3] in that it offered the panel the benefit of "quasi-anonymity" rather than total anonymity; there were very few "rules" or known scientific evidence on which to base a decision; and the panel consisted of a heterogeneous group of participants of whom the majority were selected for their decision-making abilities and clinical knowledge rather than their expertise or knowledge of automated PCISs.

Background

The Delphi technique is a process that facilitates consensus building and informed decision making among experts in a field. It is one of several group techniques developed for situations where individual judgments must be combined to arrive at informed decisions that cannot be made by one person and for which there is insufficient scientific information or an overload of often contradictory information. Four characteristics, in combination, distinguish the Delphi technique from other group decision-making processes: anonymity, iteration with controlled feedback, statistical group response, and "expert" input.[4] Although first used as a technologic forecasting process at the RAND Corporation, the Delphi technique may be used when one or more of the following occurs: a problem cannot be solved by analytic technique alone, but requires subjective judgment on a collective basis; the contributing individuals represent a diversity of experience and expertise; frequent meetings are not feasible; disagreements among individuals are so severe or politically charged that anonymity must be ensured; and the effectiveness of face-to-face meetings can be increased by a supplemental group process, i.e., avoidance of the "bandwagon" or "groupthink" effect.[5] Central to the use of this technique is the lack of agreement or incomplete state of knowledge concerning either the nature of the problem or the components that must be included in a successful solution. Group consensus is achieved through the administration of successive rounds of questionnaires in which opinions are gathered. Members of the group independently submit their opinions through responses to the questionnaires. The initial round is open-ended or exploratory and is designed to obtain information about an issue. Questionnaires for successive rounds incorporate opinions from the previous round and statistical indicators such as the mean or standard deviations of previous responses. As the rounds progress, the opinions of the group begin to merge toward consensus. Consensus is achieved when there is a clear indication of the group's opinion, generally reflected by a previously defined statistical indicator.

The Delphi technique was not used in the health care field until the mid-1970s, when early studies concentrated on determining priorities for nursing research[6] and developing a nursing research curriculum.[7] Since then, the use of the Delphi process in the clinical setting has steadily increased. The technique has been progressively used in the health care field and in particular, during the 1990s, to develop educational curricula for patients and health professionals[8]; develop and assess quality-of-care indicators[9–15]; determine criteria and priorities for standards, new programs, and research[16–25]; assess priorities, needs, and guidelines for automation in the health care field[26–29]; and develop criteria and priorities for administration.[30] (The project report for the hospital contains a brief description of each of these studies.[31])

Modifications or variations of the conventional Delphi approach are described in the literature—the policy Delphi,[32] the decision Delphi,[3] the reactive Delphi,[33] and the real-time Delphi.[5] Unlike the conventional design, the decision Delphi technique does not focus on facts or the use of specific application experts. The outcome is focused on decision making in fields that are strongly susceptible to change and where one or more the following occurs: there is more influence by individual decision makers than by underlying rules; the field of interest is relatively new or is presented with new developments; or the scientific field of interest is small and relatively self-contained.[3] The panel would include a high percentage of decision makers in the field under consideration. This differs from the conventional approach in that these people may not necessarily be specific application experts. The practical application of the decision Delphi is similar to the conventional approach. Unlike the conventional approach, a situation of "quasi-anonymity" exists, as the participants are known but their statements and comments remain anonymous.

Advantages and Limitations of the Delphi Technique

The major strength of the Delphi technique is that it can be used in a diversity of applications while providing for consensus of group opinion in a way that eliminates the negative aspects of face-to-face meetings. Whitman[34] has outlined several significant advantages of the technique over the committee meeting. Although it does not guarantee more participation than in a face-to-face meeting, it encourages honest opinion, free from peer pressure, which serves to reduce or eliminate the "bandwagon" or "groupthink" effect that can bias committee meetings or group discussions. Compared with committee meetings, the Delphi technique also has time management benefits because it eliminates "off the subject" ideas. Greater acceptance of decisions because of the ability to include large numbers of the community in the decision making, or "grassroots" involvement, has been noted.[35]

The length of the process can be a limiting factor on the respondent's motivation. It is suggested that three or four rounds are enough for people to react to the ideas of others and thus avoid the "fatigue factor" and the subsequent tendency to end the process by conforming to the group opinions.[36] The definition of an expert[37] and the selection of the group or panel respondents remain subject to considerable debate. It has been suggested that, for health care issues, to avoid getting a false consensus due to a limited range of viewpoints, a heterogeneous panel of experienced personnel is required[38] as well as predefined selection criteria for the panel members.[39] In addition, a lack of synergy can develop during a Delphi round meeting if proper procedure is not followed.

Decisions made by use of the Delphi technique are based on opinions and not necessarily on facts obtained from a controlled environment. However, when a collective decision is required from members of varied experience and expertise and when the information available on the issue is limited, the Delphi process has been shown to be of value in assessing the priorities, needs, and guidelines for automation.[26,27,29]

Methods

A modified Delphi technique was chosen primarily because the PCIS automation recommendation would affect the entire caregiver community and therefore required a collective group opinion from members with diverse perspectives, experience, and expertise.

The final outcome was the achievement of consensus on the goals and criteria, and their respective weights, each weight being equivalent to the mean score of the round when consensus was achieved. After the Delphi rounds were completed, the software products were further evaluated using these weighted criteria to arrive at the recommendation to purchase.

The Delphi Question

The question answered from this Delphi study was "What are the relevant goals and criteria and their respective weights that should be used in selecting an integrated, computerized patient care information system?"

The Duration

The modified Delphi approach was used in the latter part of the six-month evaluation process. The duration of the Delphi process, approximately four weeks, started with an open-ended questionnaire to solicit input for the criteria and ended with a two-day retreat to finalize the goals and criteria. The software recommendation was then decided using the weighted goals and criteria. A two-day retreat was chosen primarily because of the direction from senior management to arrive at the recommendation within a relatively short time. Having the participants together in a retreat setting for a specified length of time ensured that there would be no delay in reaching the decision.

The Participants

The institution's executive management appointed the chair of the selection task force, a physician respected by the administrative and medical communities. The chair then selected other caregivers as participants for the panel on the basis of the following criteria:

- They were recognized as decision makers or opinion makers in their professions,

- They were actively involved in providing patient care or in supporting patient care, including documentation and distribution of information;

- They were knowledgeable about the use of computerized systems in the clinical setting; and

- They were willing to participate for the duration of the entire six-month selection process.

The Delphi panel consisted of ten representatives from the clinical and the administrative areas of the hospital. The number of panel members was kept relatively small (for a Delphi process), as each panel

Table 1 ■

Guidelines for Evaluation

1. The system must be designed for direct use by caregivers.

2. The product must be proven in an institution similar to this hopsital and the system must be provided by a stable vendor.

3. The system must be one that can serve as a basis for re-engineering the hospital's work processes.

4. The vendor must offer a fully integrated product.

5. The selected system must provide a financial payback.

6. The system must provide the basis for a computer-based patient record.

member was expected to participate fully in at least two client site visits in addition to the on-site evaluation process. This responsibility represented a significant amount of time away from clinical duties, and to ensure that full, consistent participation would occur over the entire length of the selection process, the number of panel members was kept lower, but it remained high enough to represent the range of caregiver services across the hospital. There were four physicians (including the chair) on the panel—an orthopedic surgeon, a respirologist, a thoracic surgeon, and an oncology medicine specialist. Two patient services managers with nursing backgrounds were on the panel, one from the intensive care unit and one from the surgical units. The remaining members of the task force, appointed by senior management, included an information systems analyst with a nursing background, the director of Information Systems, the manager of Patient Care Information Systems, and the director of Patient Information and Records Management. The monitor of the study (a nonvoting member of the panel) was an information systems analyst who was responsible primarily for distributing the questionnaires, tabulating the results, assisting the chair with administrative functions, and overseeing the process. The monitor and the chair of the task force were responsible for creating and revising the questionnaires, distributing them, and addressing any issues or concerns raised during the process.

Measurement of Consensus

The panel agreed beforehand that the mean and standard deviation would be used as the primary indicators of group consensus. For weighting the goals, a standard deviation value of two units was adopted as the measure of achieving consensus. For the development of the criteria and weights, a standard deviation of one unit was adopted as the ideal consensus or agreement measure, because the range of possible

scores for each criterion was from 0 to 10, so large variations in responses were not expected. The group decided to deal with exceptions or values outside this range on an "as needed" basis during the process.

Round 1

The first questionnaire was open-ended and asked the question "What are the important criteria that should be used to make a decision on an automated patient care information system?" Prior to distribution of this question to the panel, the panel reviewed five ground rules or guidelines that were passed to them by the institution's executive management (Guidelines 1 to 5, Table 1). The group retained these five ground rules as the goals on which to develop criteria, with the addition of Guideline 6. These guidelines were subsequently referred to as the selection goals.

In addition to the selection goals, the panel members were given a sample list of criteria that had been gathered by information system analysts as part of the "preselection" phase of the project. These sample criteria, with each member's experiences gained on client site visits during the previous three months, served as the background information for the group to address the Delphi question. Each member was then asked to submit, in writing, a list of five to ten criteria for each identified goal. Three weeks were allowed for the responses. The chair and the monitor received responses from all panel members. The chair and the monitor then organized these criteria into the goal categories.

Round 2

The second and subsequent rounds of this Delphi process took place during a two-day retreat attended by all ten panel members. The Delphi process was used

Goal Statement	Weight
1. Designed for direct use by caregivers	
2. Proven product/stable vendor/future directions	
3. Basis for re-engineering work processes	
4. Fully integrated product	
5. Financial payback	
6. Basis for electronic patient record	
Total must equal:	100

Figure 2 Member's goal ballot for Round 2 of the Delphi process.

Table 2 ■

Goals and Assigned Weights (Rounds 2 and 3)

Selection Goal	Round 2		Round 3	
	Mean	SD	Mean	SD
1 Designed for direct use by caregivers	31	8	31	3.3481
2 Basis for electronic patient record	24	5.6789	15	3.7802
3 Proven product/stable vendor/future directions	15	4.4721	10	2.3324
4 Fully integrated product	11	6.6332	11	5.1196
5 Basis for re-engineering work processes	10	3.2012	9	3.0000
6 Financial payback	9	3.7417	24	2.9816
	100		100	

NOTE: The means for round 3 equal the assigned goal weights for each of the six goals. SD indicates standard deviation.

with the nominal group technique. No verbal interaction as to the matter at hand was permitted between or during the completion of the questionnaires, and each member's responses remained unknown to the rest of the group. The second questionnaire (Figure 2) was distributed to each member with written instructions for completion. These instructions required the participant to weight each goal so that the total weighting of the six goals equaled 100. The group was given 30 minutes to prepare their responses. Responses were received from all members, and the mean and standard deviation were tabulated on site using a personal computer spreadsheet software application.

Round 3

The questionnaire or goal ballot was distributed to the group with the mean and standard deviation of the group's response to Round 2 (Table 2). The panel was instructed to review this response and weight the goals using the new information within 30 minutes. The mean and standard deviation of this round's responses were then tabulated. The group reviewed the findings and decided that, on the basis of the small variation of the means, no further agreement could be achieved by another round. The means of the third-round responses for each goal then became the assigned weights for each goal (Table 2).

Round 4

To assign weights to the criteria identified in Round 1, a criterion ballot was distributed for each goal. The criterion ballot listed each goal followed by 1 to 13 criteria, as shown on Table 3. The members were asked to weight each criterion on a scale of 0 to 10, with 0 indicating the criterion is absent; 5, criterion is met adequately; and 10, criterion is met at the highest level. The total of the criterion weights did not have to equal the respective goal weight. The members

were given 60 minutes to complete the weighting. The results from this weighting were tabulated, and the mean and standard deviation for all criteria in each of the six goal categories were put onto the next round's ballot for the members' consideration.

Round 5

The next questionnaire, or second criterion ballot, was distributed to the group with the mean and standard deviation of the group's response to the first weighting of the criteria. The panel was instructed to review this response and weight the goals on the basis of this new information. Based on the predefined measure of group consensus (i.e., one standard deviation around the mean) and response distribution, consensus for the criteria of two goals—Goal 5, "basis for re-engineering," and Goal 6, "financial payback"—was recognized. The panel was given 60 minutes to prepare their responses for weighting the remaining criteria. The mean and standard deviation of the responses to the fifth-round ballot responses were then tabulated. The group reviewed the findings and decided that consensus had been attained for all criteria except three in Goal 3, "proven product/stable vendor/future directions."

Round 6

The members were given a third criterion ballot with the three criteria from Goal 3, "proven product/stable vendor/future directions" and the group's response, i.e., the mean and standard deviation, from Round 5. The members were asked to weight these criteria. The results were tabulated and reviewed by the group. There was little variation from the mean and standard deviation of the previous balloting, and the members decided that another round would not increase agreement. The assignment of the final criteria weights for all goals, shown in Table 3, was considered complete.

Table 3 ■

Goals, Criteria, and Maximum Assigned Criterion Weights

	Maximum Assigned Criterion Weight	Maximum Goal Weight
1. Designed for direct use by caregivers:		31
Supports on-line access to data from all previous episodes of care	9.33	
System availability	9.22	
Supports clinical decision-making (alerts, online reference etc.)	9.11	
Easy movement within & between patient records	8.67	
Transaction time (quick response time)	8.44	
Contains features that improve caregiver productivity	8.22	
Provides on-line Help at screen & field level	7.56	
Attractive user interface	7.33	
Number of screen flips per transaction (minimal)	7.33	
Supports free text entry of data	6.56	
Ease of creation of department/personal order sets	5.89	
Flexible & personalized views of data, including graphical display	5.78	
Computer-based training available	4.67	
2. Basis for electronic patient record:		24
Supports ad hoc user access to clinical data	9.88	
Provides unique encounter identifier	9.38	
Can merge records	9.00	
Security of system	8.50	
Appointments: strength of scheduling function	8.38	
Usability of system management tools	8.25	
3. Proven product/stable vendor/future directions:		15
Confidence in vendor's ability to provide ongoing support	9.22	
Confidence in vendor's ability to provide implementation support	8.89	
Futures: Technology and Development	8.78	
Willingness to adapt the product quickly	8.33	
Comprehensiveness of hospital starter set	7.33	
Prospects for joint product development	6.94	
Strong and active user group	4.56	
Total number of installations	2.89	
Number of sites installed in Canada	2.33	
4. Fully integrated product:		11
Provides tools to create specialty databases with main database	9.00	
Emergency Room functionality available	8.67	
Availability of integrated ancillary subsystems	7.22	
5. Basis for re-engineering work processes:		10
Electronic patient record	9.67	
Supports physician-driven order communication	8.89	
Supports multiple resource scheduling	7.11	
6. Financial payback:		9
Provides timely alerts/advice that support cost containment	9.33	
Product maximum weighted score:		100

Post-Delphi: Selecting the Product of Choice

To make a final decision as to the product-of-choice, the group then evaluated the two software products by scoring each software product against each criterion, using a scale of 0 to 10, with 0 indicating the criterion is absent; 5, criterion is met adequately; and 10, criterion is met at the highest level. To achieve the product criterion score, the mean of the group's raw scores was multiplied by the assigned criterion weight. Then the product criterion scores were totaled by goal, as were the maximum criterion scores (the maximum criterion score being the highest possible score—i.e., 10—multiplied by the assigned criterion weight). The total product criterion score was divided by the total maximum criterion score for each goal

and multiplied by the respective assigned (or maximum) goal weight score to give the product's goal weight score. For each product, the six product goal weight scores were summed to provide the overall product score. The product of choice had the highest overall product score. The scores for the two products are shown in Tables 4 and 5.

Results

Ten panel members participated for the majority of the duration of the two-day decision-making retreat. For Rounds 2 and 3, the goal determination rounds, ten responses were received. For the criteria determination rounds, Rounds 4 through 6, the number of responses varied from eight to ten, as some participants either chose not to respond or were not present for the entire process.

Goal Determination Rounds (Rounds 2 and 3)

Consensus was based on the changes in the standard deviation or on a group decision that, given the response distribution, no further gains in agreement would be achieved by another round. After the study, the results of these rounds were analyzed statistically for agreement using two methods of analysis. The first agreement analysis was the test of classic agreement: "Was there change in an *individual's* score from Round 2 to Round 3?" The group response consisted of ten paired responses ($n = 10$) where the responses of the second round represented the "before" intervention and the third round the "after," the intervention being the additional new information, i.e., the mean of the group's previous response. Bland and Altman's[40] "limits of agreement" statistical indicator and repeatability coefficient were selected for this test rather than a Student *t*-test of significance or a correlation test, because the *t*-test is irrelevant to the question of agreement. It would be very unusual if two methods designed to measure the same quantity were not related, and even data in poor agreement can produce quite high correlation. The limits of agreement are determined by comparing the differences in the means and standard deviations of each pair of individual responses. The simple repeatability coefficient, or *t*-value, is determined by squaring the differences, summing the squared differences, dividing by *n* (or number of paired responses), and taking the square root to get the standard deviation of the differences.[4] The expectation is that 95 percent of differences will be less than two standard deviations (assuming a normal distribution and a mean difference of zero). The alpha value used was 0.05 (assuming a two-tailed test)

with degrees of freedom of *n* minus 1, or 9. There was no significant change in agreement of any individual's responses for any goal from Round 2 to 3. It could be concluded that agreement of individual responses between both goal determination rounds was strong. However, no conclusions can be made as to the effect of the new information on the responses or the tendency to conform to the mean of the previous response.

The second agreement analysis was a test of difference in variability of the group's response in Round 2 compared with Round 3: "Was there a change in *group* consensus from Round 2 to 3?" A one-sided *F*-test for two sample variances was used, with an alpha value of 0.05. A one-sided test was used since it was assumed that variation in agreement would either remain the same or be reduced, which is a valid assumption for a consensus exercise. The results showed a significant movement toward consensus from Round 2 to Round 3 responses for Goal 1, Designed for direct use by caregivers ($P = 0.008$), and for Goal 5, Basis for re-engineering work processes ($P = 0.0329$). Otherwise, for the remaining goals, there was no significant change, implying that there was already a sufficient level of consensus. The results of the statistical analysis supported the decisions made by the group that consensus for goal determination was achieved by the end of Round 3.

Criteria Determination Rounds (Rounds 4 to 6)

The criteria for all goals, with one exception, were assigned final weights after Rounds 4 and 5. The sixth round (or third criterion weighting iteration) was necessary to resolve the dissensus over several criteria for Goal 3, "proven product/stable vendor/future directions." The group used the standard deviation of the previous round as the guideline to decide on the achievement of consensus. The two tests of agreement, the test for classic agreement, and the test of difference in group consensus were not done for the criteria results because the range of response scores (0 to 10) was very small, so any significant variation would have been surprising. Also, for some criteria there was not a "paired" individual response, as the responses varied in number from 8 to 10 on a given round. Only one round (Round 4) of scoring was required for Goals 5 and 6, as the distribution of responses indicated a relatively high level of agreement, with less than two standard deviations about the mean. By the end of the fifth round, there was sufficient reduction in the standard deviation for the group to agree that no further agreement would be reached by another round. The mean scores of each

Table 4 ■

Calculation of Overall Product Score for Product A

	Mean of Consensus Round	Criterion Weight	Product Criterion Score	Maximum Criterion Score	Maximum Goal Weight Score	Product Goal Weight Score*
1. Designed for direct use by caregivers:					31	25.2
Supports online access to data from all previous episodes of care	9.50	9.33	88.64	93.30		
System availability	9.50	9.22	87.61	92.22		
Supports clinical decision making (alerts, online reference, etc.)	8.70	9.11	79.26	91.10		
Easy movement within and between patient records	9.10	8.67	78.86	86.66		
Transaction time (quick response time)	6.67	8.44	56.32	84.44		
Contains features that improve caregiver productivity	9.30	8.22	76.46	82.22		
Provides online Help at screen and field level	9.30	7.56	70.27	75.56		
Attractive user interface	8.80	7.33	64.53	73.33		
Number of screen flips per transaction is minimal	8.89	7.33	65.19	73.33		
Supports free text entry of data	7.90	6.56	51.78	65.55		
Ease of creation of department/personal order sets	5.90	5.89	34.74	58.88		
Flexible and personalized views of data, including graphic display	7.00	5.78	40.44	57.77		
Computer-based training available	0.56	4.67	2.61	46.66		
			796.71	981.02		
2. Basis for electronic patient record:					24	19.6
Supports ad hoc user access to clinical data	9.1	9.88	89.86	98.75		
Provides unique encounter identifier	8.7	9.38	81.56	93.75		
Can merge records	9.6	9.00	86.40	90.00		
Security of system	6.56	8.50	55.76	85.00		
Appointments: strength of scheduling function	6	8.38	50.25	83.75		
Usability of system management tools	8.78	8.25	72.44	82.50		
			436.27	533.75		
3. Proven product/stable vendor/future directions:					15	10.8
Confidence in vendor's ability to provide ongoing support	8.2	9.22	75.62	92.22		
Confidence in vendor's ability to provide implementation support	7.2	8.89	63.99	88.88		
Futures: technology and development	9.3	8.78	81.65	87.80		
Willingness to adapt the product quickly	8.3	8.33	69.16	83.33		
Comprehensiveness of hospital starter set	6.5	7.33	47.66	73.33		
Prospects for joint product development	8	6.94	55.55	69.44		
Strong and active user group	4.6	4.555	20.95	45.55		
Total number of installations	3.6	2.888	10.40	28.88		
Number of sites installed in Canada	0.8	2.333	1.87	23.33		
			426.86	592.76		
4. Fully integrated product:					11	9.3
Provides tools to create specialty databases with main database	9	9	81.00	90.00		
Emergency room functionality available	7.4	8.667	64.14	86.67		
Availability of integrated ancillary subsystems	9	7.222	65.00	72.22		
			210.13	248.89		

5. Basis for re-engineering work processes:						
Electronic patient record	9.3	9.666	89.89	96.66	10	8.4
Supports physician-driven order communication	7.5	8.888	66.66	88.88		
			156.55	185.54		
6. Financial payback:						
Provides timely alerts/advice that support cost containment	8.6	9.333	80.26	93.33	9	7.7
			80.26	93.33		
Overall product score for Product A:					100	81.1

*The product goal weight score (PGWS) was calculated as follows:

$$PGWS = (PCS/MCS) \times MGWS$$

where PCS is the product criterion score; MCS, the maximum criterion score; and MGWS, the maximum goal weight score. Thus, the PGWS for Goal 1 for Product A is: $(796.71/981.02) \times 31 = 25.2$.

criterion for the last round represented the final or assigned weight, that is, the weight that would be used to evaluate the two software products (Table 3).

Evaluation of the Approach

At a wrap-up session immediately after the decision-making retreat, the panel members indicated verbally that they were in complete agreement with the outcome of the Delphi decision-making process and with the process itself. A formal evaluation of the success of this approach was not performed. As for the soundness or success of the actual product decision, an evaluation will be completed once the PCIS is fully operational.

Discussion

The end results of the study, the weighted goals and criteria, indicate the priority of the panel to recommend a PCIS software product that focused on main source of information in the hospital, the patient record, and robustly addressed the functionality required of a frontline caregiver. This is reflected in the goal weights, the number of criteria, and the weights assigned to the goals "direct use by caregiver" (31/100) and "basis for an electronic patient record" (24/100).

Several features of this Delphi process tended to facilitate consensus. First, involvement of clinicians in the selection of a product that they will have to use in their day-to-day practice minimized the perception that an outside expert was trying to dictate their method of practice. The fact that the majority of clinical members of this selection task force agreed to remain as key members of the PCIS implementation advisory council, where their primary responsibility would be to recommend and encourage acceptance of this software product to their fellow workers, also reflected their agreement. Second, the selection of a clinician rather than an administrator to lead the panel helped minimize the perception of outsiders trying to manage clinical practice. Third, the use of successive questionnaires to gather information and outline the group's collective response of the previous round with statistical indicators promoted a sense of group ownership of the process and thus facilitated consensus building. Fourth, the anonymity of responses (although not the participants' identity) ensured that the group was not unduly influenced by one or two members of the group or by the biases of the individuals reflecting the biases inherent in the hierarchic structure of the hospital. Finally, this process allowed the focus to remain on the selection of the PCIS project while at the same time eliminating the "groupthink"

Table 5 ■

Calculation of Overall Product Score for Product B

	Mean of Consensus Round	Criterion Weight	Product Criterion Score	Maximum Criterion Score	Maximum Goal Weight Score	Product Goal Weight Score*
1. Designed for direct use by caregivers:					31	18.5
Supports online access to data from all previous episodes of care	6.20	9.33	57.85	93.30		
System availability	6.40	9.22	59.02	92.22		
Supports clinical decision making (alerts, online reference, etc.)	7.80	9.11	71.06	91.10		
Easy movement within and between patient records	5.50	8.67	47.66	86.66		
Transaction time (quick response time)	7.00	8.44	59.11	84.44		
Contains features that improve caregiver productivity	4.50	8.22	37.00	82.22		
Provides online Help at screen and field level	5.80	7.56	43.82	75.56		
Attractive user interface	3.80	7.33	27.87	73.33		
Number of screen flips per transaction is minimal	4.00	7.33	29.33	73.33		
Supports free text entry of data	4.44	6.56	29.10	65.55		
Ease of creation of department/personal order sets	8.40	5.89	49.46	58.88		
Flexible and personalized views of data, including graphic display	5.60	5.78	32.35	57.77		
Computer-based training available	8.78	4.67	40.97	46.66		
			584.60	981.02		
2. Basis for electronic patient record:					24	13.4
Supports ad hoc user access to clinical data	2.70	9.88	26.66	98.75		
Provides unique encounter identifier	6.30	9.38	59.06	93.75		
Can merge records	3.90	9.00	35.10	90.00		
Security of system	7.56	8.50	64.26	85.00		
Appointments: strength of scheduling function	8.89	8.38	74.45	83.75		
Usability of system management tools	4.78	8.25	39.44	82.50		
			298.97	533.75		
3. Proven product/stable vendor/future directions:					15	11.1
Confidence in vendor's ability to provide ongoing support	8.80	9.22	81.15	92.22		
Confidence in vendor's ability to provide implementation support	9.40	8.89	83.55	88.88		
Futures: technology and development	6.40	8.78	56.19	87.80		
Willingness to adapt the product quickly	5.90	8.33	49.16	83.33		
Comprehensiveness of hospital starter set	8.40	7.33	61.60	73.33		
Prospects for joint product development	4.30	6.94	29.86	69.44		
Strong and active user group	8.90	4.56	40.54	45.55		
Total number of installations	7.60	2.89	21.95	28.88		
Number of sites installed in Canada	6.00	2.33	14.00	23.33		
			438.00	592.76		
4. Fully integrated product:					11	4.6
Provides tools to create specialty databases with main database	4.00	9.00	36.00	90.00		
Emergency room functionality available	4.10	8.67	35.53	86.67		
Availability of integrated ancillary subsystems	4.40	7.22	31.78	72.22		
			103.31	248.89		

5. Basis for re-engineering work processes:					
Electronic patient record	6.60	9.67	63.80	96.66	10	7.7
Supports physician-driven order communication	8.90	8.89	79.10	88.88		
			142.90	185.54		
6. Financial payback:					
Provides timely alerts/advice that support cost containment	6.00	9.33	56.00	93.33	9	5.4
			56.00	93.33		
Overall product score for Product A:					100	60.7

*The product goal weight score (PGWS) was calculated as follows:

$$PGWS = (PCS/MCS) \times MGWS$$

where PCS is the product criterion score; MCS, the maximum criterion score; and MGWS, the maximum goal weight score. Thus, the PGWS for Goal 1 for Product B is: $(584.60/981.02) \times 31 = 18.5$.

effect of regular committee meetings and opinion-gathering sessions.

Several limitations were identified. Although completing six rounds of the study and scoring the two products during a two-day retreat ensured a captive audience, and thus a high response rate, the probability of a "fatigue factor" by the end of rounds was high. Although this was not mentioned by the members, we cannot state with total confidence that the results represented group consensus based totally on informed decision making rather than an interest in completing the process, particularly by the end of the second day. Because of the need to complete the exercise in a relatively short time (for a Delphi study), it is also possible that not all issues and concerns were addressed during the exercise to every member's satisfaction in order to make an informed decision.

A question that was not answered specifically through the design methodology of this process was "How reliable were the weighted goals and criteria as determined during this process?" Although the Delphi technique is considered more an art than a science, giving the questionnaires to a control group, or a group external to the process, would have helped validate the structure, clarity, and ease-of-use of the process. For example, a pilot round on five to ten individuals outside the study before the goal determination rounds could have been incorporated into the design and the responses evaluated as in a test–retest reliability research study.[41] Another way to improve the reliability and validity of the study would be to designate a content expert to provide clarification and guidance during the process to those members considered nonexperts. For example, a consultant specializing in information systems technology would be able to address questions from the primary caregivers, and a clinical specialist would be able to address questions from the administrative membership. This would help reduce potential bias introduced by the chair if the responsibility for answering questions depended mainly on the chair. In addition, the incorporation of automated communication systems to increase the information available to participants through side conversations or conversation histories while preserving the anonymity of those asking the questions would speed up the distribution of information and would allow a participant to review more information related to a specific item or issue.[42] Having quicker access to more relevant statistical indicators besides the mean and standard deviation between rounds, specifically the group's distribution trends and the tests of agreement, might also reduce the number of rounds required and thus lessen

or eliminate the potential fatigue factor associated with too many rounds.

The responses of the participants should have been evaluated through completion of an attitudinal questionnaire shortly after the decision-making retreat to evaluate the benefits of using the Delphi approach. Such a formal evaluation would have been useful in determining the construct validity of this process. This type of formal evaluation was not possible for several reasons. Four of the clinical panel members left the institution before the implementation to pursue other career opportunities, and with the majority of the remaining panel members taking on leadership roles for the implementation, it was thought that their response to such an evaluation would be biased. Evaluation of the usefulness of this approach was based then mainly on informal, verbal comments of the panel members made immediately after the retreat ended.

In spite of the limitations, the PCIS selection task force did develop, in a short time, the goals and criteria in a manner that indicated strong agreement among members of a rather diverse group. Also, the design of this Delphi study helped increase the validity of the study in two ways: first, the goals and criteria were developed by the people who will actually use the PCIS, so the goals and criteria have a high face validity; and second, concurrent validity was achieved as consensus was achieved among the experts themselves.

Conclusion

The relevant goals and criteria determined through this process and used to make the final decision for software selection were developed as a result of strong agreement among members of a diverse group in a relatively short time. A consensus-building process such as the Delphi technique was indicated in this setting because of the need to select one PCIS that would be used by all caregivers and administrative staff in the institution while avoiding a decision based on political biases and the "groupthink" effect of traditional committee meetings. It is recognized that the Delphi technique is a semiquantitative, semiqualitative research method, in that information and a decision are based on opinions and not necessarily on facts obtained from a controlled environment. However, since the PCIS selection task force was representative of the prospective users of such a system, since there was little known information at the outset, and since the decision was made within the required time frame, it is concluded that the Delphi process was use-

ful in determining the goals and criteria to use in the selection of an automated PCIS at Vancouver Hospital and Health Sciences Centre.

References ■

1. Janis IL. Groupthink: Psychological Studies of Policy Decisions and Fiascoes. Boston, Mass.: Houghton Mifflin, 1982: 9.
2. Dalkey N, Helmer O. An experimental application of the Delphi method to the use of experts. Manage Sci. 1963;9: 458–67.
3. Rauch W. The decision Delphi. Technol Forecast Soc Change. 1979;15:159–69.
4. Goodman CM. The Delphi technique: a critique. J Adv Nurs. 1987;12:729–34.
5. Linstone HA, Turoff M. The Delphi Method: Techniques and Applications. Reading, Mass.: Addison Wesley, 1975.
6. Lindeman CA. Delphi survey of priorities in clinical nursing research. Nurs Res. 1975;24:434–41.
7. Sullivan E, Brye C. Nursing's future: use of the Delphi technique for curriculum planning. J Nurs Educ. 1983;22:187–9.
8. McGoldrick TB, Jablonski RS, Wolf ZR. Needs assessment for a patient education program in a nursing department: a Delphi approach. J Nurs Staff Dev. 1994;10:123–30.
9. Ashton C, Kuykendall DH, Johnson ML, et al. A method of developing and weighting explicit process of care criteria for quality assessment. Med Care. 1994;32:755–70.
10. Hutchinson A, Fowler P. Outcome measures for primary health care: what are the research priorities? Br J Gen Pract. 1992;42:227–331.
11. Kitson A, Harvey G, Hyndman S, Yerrell P. A comparison of expert- and practitioner-derived criteria for post-operative pain management. J Adv Nurs. 1993;18:218–32.
12. Lazaro P, Fitch K. From universalism to selectivity: is appropriateness the answer? Health Policy. 1996;36:261–72.
13. McDonnell J, Meijler A, Kahan JP, Bernstein SJ, Rigter H. Panelist consistency in the assessment of medical appropriateness. Health Policy. 1996;37:139–52.
14. Megel ME, Barna ME, Rausch AK. Conflicts experienced by quality assurance/improvement professionals: a Delphi study. J Nurs Care Qual. 1996;10:75–82.
15. Moussa A, Bridges-Webb C. Quality of care in general practice: a Delphi study of indicators and methods. Aust Fam Physician. 1994;23:465–73.
16. Broome ME, Woodring BW, O'Connor-Von S. Research priorities for the nursing of children and their families: a Delphi study. J Pediatr Nurs. 1996;11:281–7.
17. Gallagher M, Bradshaw C, Nattress H. Policy priorities in diabetes care: a Delphi study. Qual Health Care. 1996;5:3–8.
18. Gruber M. The Development of a Position Statement Using the Delphi Technique. Gastroenterol Nurs. 1993;16:68–71.
19. Harrington JM, Calvert IA. Research priorities in occupa-

tional medicine: a survey of United Kingdom personnel managers. Occup Environ Med. 1996;53:642–4.

20. Jairath N, Weinstein J. The Delphi methodology, part 2: a useful administrative approach. Can J Nurs Admin. 1994;7: 7–20.

21. Ketelaars C, Saad H, Halfens RJG, Wouters EF. Process standards of nursing care for patients with COPD: validation of standards and criteria by the Delphi technique. J Nurs Care Qual. 1994;9:78–86.

22. Naylor CD, Williams JI, for the Ontario Panel on Hip and Knee Arthroplasty. Primary hip and knee replacement surgery: Ontario criteria for case selection and surgical priority. Qual Health Care. 1996;5:20–30.

23. Panniers TL, Walker EK. A decision-analytic approach to clinical nursing. Nurs Res. 1994;43:245–9.

24. Pearson SD, Margolis CZ, Davis S, Schreier LK, Sokol HN, Gottlieb LK. Is consensus reproducible: a study of an algorithmic guidelines development process. Med Care. 1995; 33:643–60.

25. Williard RL, Tresolini CP, O'Neil EH. Characteristics, importance, and implications of comprehensive drug therapy management. Am J Health Syst Pharmacists. 1996;53:623–32.

26. Carter BE, Axford RL. Assessment of computer learning needs and priorities of registered nurses practicing in hospitals. Comput Nurs. 1993;11:122–6.

27. Green CG, Khan MA, Badinelli R. Use of the Delphi research technique to test a decision model in foodservice systems: a case study in food production. J Am Diet Assoc. 1993;93:1307–9.

28. Lobach DF. A model for adapting clinical guidelines for electronic implementation in primary care. Proc. 19th Annu Symp Comput Appl Med Care. 1995:581–5.

29. Weir C, Lincoln M, Roscoe D, Turner C, Moreshead G. Dimensions associated with successful implementation of a hospital-based integrated order entry system. Proc 18th Annu Symp Comput Appl Med Care. 1994:653–7.

30. Hudak RP, Brooke PP, Finstuen K. Forecast 2000: a prediction of skill, knowledge, and abilities required by senior medical treatment facility leaders into the 21st century. Milit Med. 1994;159:494–500.

31. Chocholik JK. A review of the Delphi technique and its application in the Patient Care Information System Selection Project at Vancouver Hospital and Health Sciences Centre [unpublished project report] Vancouver, British Columbia, Canada: Vancouver Hospital and Health Sciences Centre, Aug 25, 1997:37–44.

32. Jairath N, Weinstein J. The Delphi methodology, part 1: a useful administrative approach. Can J Nurs Admin. 1994;7: 29–42.

33. McKenna HP. The Delphi technique: a worthwhile research approach for nursing? J Adv Nurs. 1994;19:1221–5.

34. Whitman NI. The Delphi technique as an alternative for committee meetings. J Nurs Admin. 1990;29:377–9.

35. McKenna HP. The Delphi technique: a worthwhile research approach for nursing? J Adv Nurs. 1994;19:1221–5.

36. Whitman NI. The committee meeting alternative: using the Delphi technique. J Nurs Admin. 1990;20:30–6.

37. Goldschmidt PG. Scientific inquiry or political critique? Technol Forecast Soc Change. 1975;7:195–213.

38. Synowiez BB, Synowiez PM. Delphi forecasting as a planning tool. Nurs Manage. 1990;21:18–9.

39. Williams PL, Webb C. The Delphi technique: a methodological discussion. J Adv Nurs. 1994;19:180–6.

40. Bland JM, Altman DG. Statistical methods for assessing agreement between two methods of clinical measurement. Lancet. 1986;307–10.

41. Rudy SF. A review of Delphi surveys conducted to establish research priorities by specialty nursing organization from 1985 to 1995. ORL Head Neck Nurs. 1996;14:16–24.

42. Adler M, Ziglio E. Gazing into the Oracle: The Delphi Method and Its Application to Social Policy and Public Health. London, England: Jessica Kingsley Publishers, 1996.